n/a
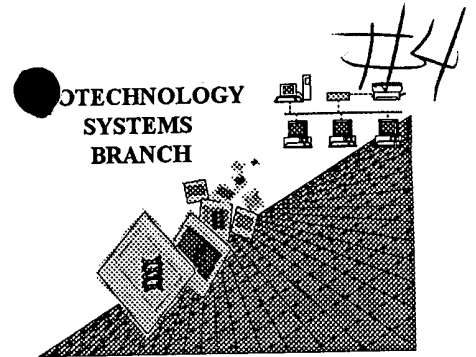
/●3

CRF Diskette Problem Report

The Scientific and Technical Information Center (STIC) experienced a problem when processing the following CRF diskette:

Application Serial Number: 09/423,093

Filing Date: 11/1/99

Classification: 536

Date Processed by STIC: 3/29/2000

Point-of-Contact/Telephone No.: __Mark Spencer 703 - 308 - 4212__

Nature of Problem:

The diskette was:

__ Damaged

__ Unreadable

__ Blank (no files present on diskette)

__ Contained an empty file (filename present, but no bytes in file)

__ A computer virus was detected on the diskette. The STIC will not process the diskette through the Data Capture System.

    Name of virus: _____

__ The Sequence Listing was not saved in ASCII (DOS) text.

✓ The diskette contained an error that disrupted normal processing:
Four separate files, representing Sequence 1-4, were submitted on disk (see attached), contrary per Sequence Rules.

```
A:\>DIR

Volume in drive A has no label
Directory of A:\

SEQ-ID~1 3          12,649   10-27-99 10:23a  Seq-ID-No.3
SEQ-ID~1 1          14,758   10-27-99 10:22a  Seq-ID-No.1
SEQ-ID~1 2          14,258   10-27-99 10:21a  Seq-ID-No.2
SEQ-ID~1 4          22,449   10-27-99 10:24a  Seq-ID-No.4
           4 file(s)          64,114 bytes
           0 dir(s)        1,361,920 bytes free
```

Files, on submitted disk, for each sequence

Per 1.824(a)(1) of Sequence Rules, the
"computer readable form shall contain a single
'Sequence Listing' ..."

Do NOT submit separate sequence number
files. ALL files must be shown in <u>one</u>
Sequence Listing. See attached.

```
GATCTGATGGCCGTAGGGCGCTACGTGCTTTCTGCTGATATCTGGGCTGAGTTGGAAAAA
ACTGCTCCAGGTGCCTGGGGACGTATTCAACTGACTGATGCTATTGCAGAGTTGGCTAAA
AAACAGTCTGTTGATGCCATGCTGATGACCGGCGACAGCTACGACTGCGGTAAGAAGATG
GGCTATATGCAGGCATTCGTTAAGTATGGGCTGCGCAACCTTAAAGAAGGGGCGAAGTTC
CGTAAGAGCATCAAGAAGCTACTGAGTGAGTAGAGATTTACACGTCTTTGTGACGATAAG
CCAGAAAAAATAGCGGCAGTTAACATCCAGGCTTCTATGCTTTAAGCAATGGAATGTTAC
TGCCGTTTTTTATGAAAAATGACCAATAATAACAAGTTAACCTACCAAGTTTAATCTGCT
TTTTGTTGGATTTTTTCTTGTTTCTGGTCGCATTTGGTAAGACAATTAGCGTGAGTTTTA
GAGAGTTTTGCGGGATCTCGCGGAACTGCTCACATCTTTGGCATTTAGTTAGTGCACTGG
TAGCTGTTAAGCCAGGGGCGGTAGCTTGCCTAATTAATTTTTAACGTATACATTTATTCT
TGCCGCTTATAGCAAATAAAGTCAATCGGATTAAACTTCTTTTCCATTAGGTAAAAGAGT
```

↑

Excerpt from one of the files on submitted disk — invalid format

The above is in _invalid format_, per Sequence Rules.

Please: 1) Consult new Sequence Rules
2) see attached sample Sequence Listing, for valid format

Appendix A To Subpart G to Part 1—Sample Sequence Listing

<110> Smith, John

Smith, Jane

<120> Example of a Sequence Listing

<130> 01-00001

<140> US 08/999,999

<141> 1998-02-28

<150> EP 91000000

<151> 1997-12-31

Please consult

<160> 2

<170> PatentIn ver. 2.0

<210> 1

<211> 403

<212> DNA

<213> Paramecium aurelia

<220>

<221> CDS

<222> 341..394

<300>

<301> Doe, Richard

<302> Isolation and Characterization of a Gene Encoding a

Protease from Paramecium sp.

<303> Journal of Fictional Genes

<304> 1

<305> 4

<306> 1 - 7

<307> 1988-06-20

<400> 1

ctactctact ctactctcat ctactatctt ctttggatct ctgagtctgc ctgagtggta   60

ctcttgagtc ctggagatct ctcctctcac atgtgatcgt cgagactgac cgatagatcg   120

ctgactgact ctgagatagt cgagcccgta cgagacccgt cgagggtgac agagagtggg   180

cgcgtgcgcg cagagcgccg cgccggtgcg cgcgcgagtg cgcggtgggc cgcgcgaggg   240

ctttcgcggc agcggcggcg ctttccggcg cgcgcccgtc cgcccctaga cctgagaggt   300

cttctcttcc ctcctcttca ctagagaggt ctatatatac atg gtt tca atg ttc   355

                                                 Met Val Ser Met Phe

```
agc ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgtttgctc        403
```

```
Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val
                   10                    15
```

<210> 2

<211> 18

<212> PRT

<213> Paramecium aurelia

<400> 2

```
Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu

  1              5                10               15


     Phe Val
```

identifiers and their accompanying information as shown in the following
table. The numeric identifier shall be used only in the "Sequence
Listing." The order and presentation of the items of information in the
"Sequence Listing" shall conform to the arrangement given below. Each
item of information shall begin on a new line and shall begin with the
numeric identifier enclosed in angle brackets as shown. The submission
of those items of information designated with an "M" is mandatory. The
submission of those items of information designated with an "O" is
optional. Numeric identifiers <110> through <170> shall only be set
forth at the beginning of the "Sequence Listing." The following table
illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	'Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd.	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO:#:	Response shall be an integer representing the SEQ ID NO shown	M
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M

<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
<222>	Location	Specify location within sequence; where appropriate state number of first and last bases/amino acids	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified

		in feature	base was used in a sequence
<223>	Other Information	Other relevant information; four lines maximum	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<300>	Publication Information	Leave blank after <300>	O
<301>	Authors	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	O
<302>	Title		O
<303>	Journal		O
<304>	Volume		O
<305>	Issue		O
<306>	Pages		O
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy	O
<308>	Database Accession Number	Accession number assigned by database including database name	O
<309>	Database Entry Date	Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy	O
<310>	Patent Document Number	Document number; for patent-type citations only. Specify as, for example, US 07/999,999	O

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	O
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	O
<313>	Relevant Residues	FROM (position) TO (position)	O
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line pre-ceding the actual sequence	M

5. Section 1.824 is revised to read as follows:

   1.824 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by   1.821(e) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;